NPS55-87-016

NAVAL POSTGRADUATE SCHOOL

Monterey, California

A DEMYANOV-TYPE MODIFICATION FOR
GENERALIZED LINEAR PROGRAMMING

SIRIPHONG LAWPHONGPANICH
DONALD W. HEARN

DECEMBER 1987

Approved for public release; distribution unlimited.

Prepared for:
Chief of Naval Research
Arlington, VA 22217

FedDocs
D 208.14/2
NPS-55-87-016

NPS-55-82-006

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral R. C. Austin
Superintendent

K. T. Marshall
Acting Provost

The work reported herein was supported in part by the Foundation Research Program of the Naval Postgraduate School with funds provided by the Chief of Naval Research and the National Science Foundation.

Reproduction of all or part of this report is authorized.

This report was prepared by:

SE URITY CLASSIFICATION OF THIS PAGE

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): NPS55-87-016	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION: Naval Postgraduate School; 6b. OFFICE SYMBOL (If applicable): Code 55	7a NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State, and ZIP Code): Monterey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING / SPONSORING ORGANIZATION: Chief of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N0001487WRE011
8c ADDRESS (City, State, and ZIP Code): Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N	RR014-01	10P-040	

11. TITLE (Include Security Classification)

A DEMYANOV-TYPE MODIFICATION FOR GENERALIZED LINEAR PROGRAMMING

12. PERSONAL AUTHOR(S)
Lawphongpanich, Siriphong; Hearn, Donald W. (University of Florida)

13a. TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ________ TO ________	1987 December	28

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Linear Programming, Decomposition, Lagrangian Dual, Subgradient

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Hearn and Lawphongpanich (1987) studied the properties of the direction formed by taking the difference of two successive dual iterates of generalized linear programming (GLP), and pointed out that this direction is also a solution to an associated direction finding problem. In this study, we show that this direction finding problem belongs to a new class of direction finding problems and propose a modification of GLP in which its original direction finding problem is replaced by another in this new class. This new direction finding problem is similar to the one used by Demyanov for minimax problems and guarantees an ascent direction for the dual function. Finally, we state and prove the convergence for the modified GLP.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Siriphong Lawphongpanich	22b TELEPHONE (Include Area Code): (408)646-2106; 22c OFFICE SYMBOL: Code 55Lp

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete — SECURITY CLASSIFICATION OF THIS PAGE

☆ U.S. Government Printing Office 1986-606-

A Demyanov-Type Modification for Generalized Linear Programming

by

Siriphong Lawphongpanich*
Donald W. Hearn**

December, 1987

* Department of Operations Research
U.S. Naval Postgraduate School
Monterey, California 93943

** Department of Industrial and Systems Engineering
University of Florida
Gainesville, Florida 32611

This research was supported in part by NSF grants ECE-8420830 and ECS-8516365 and a grant from the Research Foundation, U.S. Naval Postgraduate School.

1. INTRODUCTION

Hearn and Lawphongpanich (1987) studied the ascent nature of generalized linear programming (GLP) which is also known as the dual cutting plane or primal column generation technique. In particular, they examined the properties of the direction, d_{GLP}, defined as the difference of two successive dual iterates generated by GLP with respect to the Lagrangian dual function, L. They showed that d_{GLP} is an ascent direction for L at points where L is differentiable. At nondifferentiable points, the column entering the master problem is not unique and an arbitrary choice can make d_{GLP} a nonascent direction. To obtain an additional insight on the ascent nature of d_{GLP}, they also showed that d_{GLP} is a solution to a direction finding problem for L.

In this paper, we introduce a new class of direction finding problems and showed that it includes the one which produces d_{GLP}. This class of problems contains direction finding problems which always produce ascent directions as well as those that do not, e.g., the d_{GLP} direction finding problem. In order to improve the rate of convergence of GLP, we replace the d_{GLP} direction finding problem with another from this new class. This new direction finding problem which we describe below guarantees to produce ascent directions and is similar to the one used by Demayanov for the minimax problem. Moreover, a line search step is also included in this modification of GLP. Hearn and Lawphongpanich (1987) indicated via a numerical example that a line search step reduces the number of iterations for GLP.

For the remainder of the paper, we state GLP algorithm in both the cutting plane and the column generation form and define the associated Lagrangian dual problem in Section 2. Then, Section 3 describes the new class of direction finding problems. Finally, we present the modification of GLP and its convergence analysis in Section 4.

2. PRELIMINARIES

Consider the primal problem:

$$\text{P1:} \qquad f^* = \min_{x} \quad f(x)$$

$$\text{s.t.} \quad g(x) \le b,$$

$$x \in X,$$

where f is a continuous real-valued function, g is a continuous function from R^n to R^m, X is a nonempty compact subset of R^n, and b is a vector in R^m. The Lagrangian dual of (P1) is:

$$\text{D1:} \qquad L^* = \text{maximize } \{ L(u) : u \ge 0, \ u \in R^m \}$$

where $L(u) = \text{minimum } \{f(x) + u[g(x)-b] : x \in X\}$, and xy denotes the usual dot product between vectors x and y.

When L(u) is relatively easy to evaluate, the pair of problems, (P1) and (D1), can be addressed by GLP. Below we state the algorithm in the dual cutting plane form (Zangwill, 1969).

<u>The Dual Cutting Plane Algorithm</u>

<u>Step 0</u>. Find a point $x_0 \in X$ such that $g(x_0) < b$. Let k = 1, and go to Step 1.

<u>Step 1</u>. Solve the k-th master problem:

$$\text{M1:} \qquad \max_{(w,u)} \quad w$$

$$\text{s.t.} \quad w \le f(x_i) + u[g(x_i)-b] \qquad \text{for } i = 0,\ldots,k-1$$

$$u \ge 0$$

Let (w_k, u_k) be an optimal solution and go to Step 2.

Step 2. Solve the k-th subproblem:

S1: $$\min \{ f(x) + u_k[g(x)-b] : x \in X \}$$

Let x_k be an optimal point, and let $L(u_k) = f(x_k) + u_k[g(x_k)-b]$. If $w_k = L(u_k)$, u_k is an optimal dual solution. Otherwise, if $w_k > L(u_k)$, then replace k by k + 1, and go to Step 1.

The k-th master problem (M1) is a linear programming problem with the following dual:

M2: $$\min_{\pi} \sum_{i=0}^{k-1} \pi_i f(x_i)$$

$$\text{s.t.} \sum_{i=0}^{k-1} \pi_i g(x_i) \le b$$

$$\sum_{i=0}^{k-1} \pi_i = 1,$$

$$\pi_i \ge 0 \quad \text{for } i = 0,\dots,k-1$$

When (M2) replaces (M1) in Step 1, the resulting algorithm is generally known as Dantzig-Wolfe decomposition (1960, 1961), column generation or GLP. Geoffrion (1970) also classifies the algorithm in the cutting plane form as the strategy of outer linearization and relaxation, or, in the column generating form as the strategy of inner linearization and restriction.

3. A CLASS OF DIRECTION FINDING PROBLEMS

Consider now a feasible direction scheme for solving (D1). At a given feasible point u_k and a parameter α, one possible direction finding problem is

$$\text{DF1:} \qquad \max_{v} \quad [g(x_k)-b](v-u_k)$$

$$\text{s.t.} \quad \alpha \leq L(v),$$

$$v \geq 0,$$

where x_k is a solution of (S1), so that $[g(x_k)-b]$ is a subgradient of L at u_k. The first constraint of (DF1) requires that v be in the level set of L as defined by α. If $\alpha > L(u_k)$, any feasible solution v to (DF1) would produce an ascend direction of the form $v-u_k$, and the best value for α is L^*. Figure 3.1 illustrates an instance of (DF1) in which α is set to $L(u_k)$ (= 7.0) and the optimal solution of the resulting problem is denoted by v_k. Note that the direction v_k-u_k is an ascent direction for any choice of subgradients.

In practice, L(v) must be approximated since it is not always available in closed form. Each different approximation would yield a different direction finding problem. Similarly, the different choices for α and x_k would also yield different direction finding problems. To ensure that v_k-u_k is an ascent direction, the approximating function along with the value for α and the vector x_k must be carefully chosen. Below, we examine two different direction finding problems derived from (DF1).

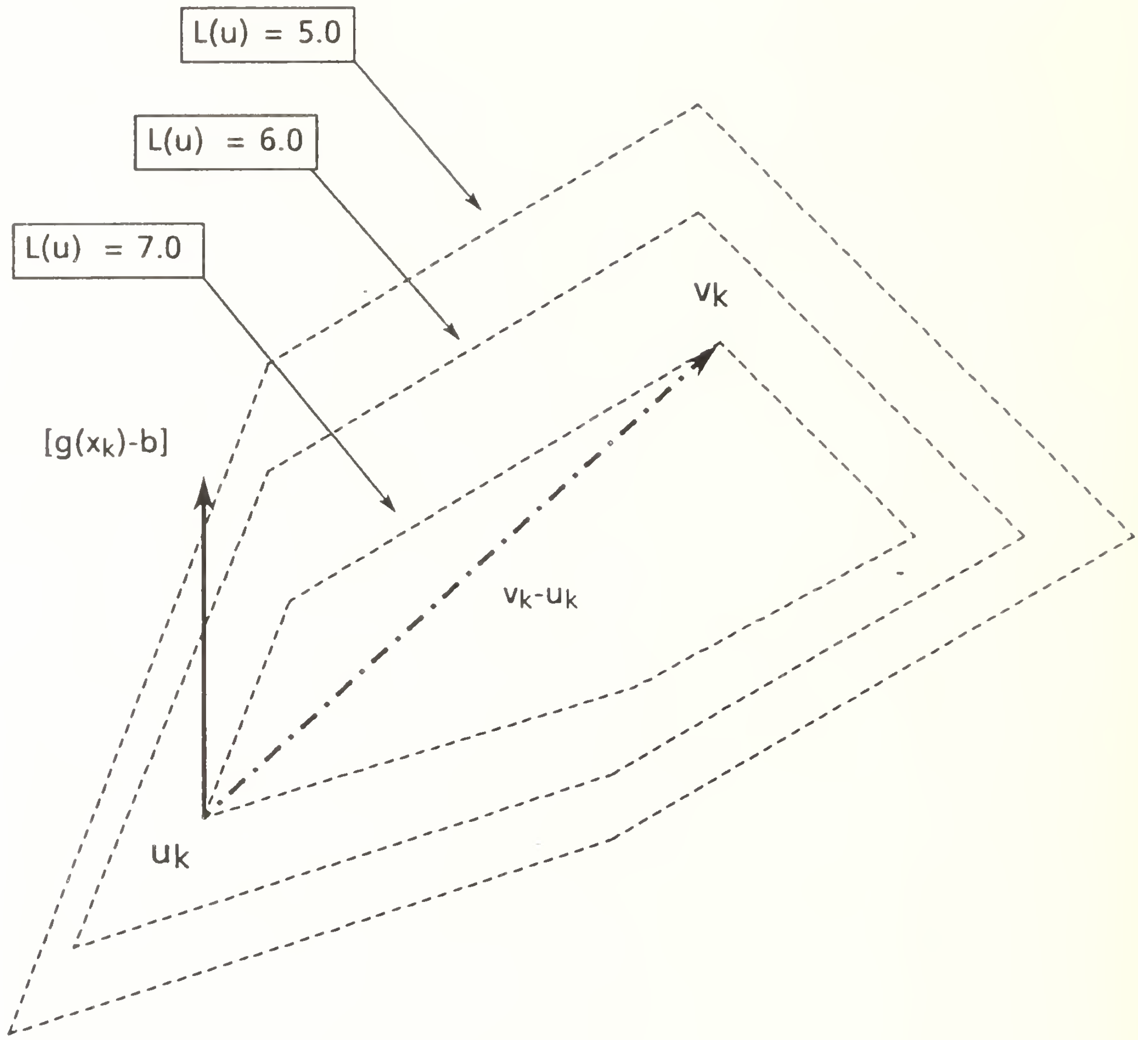

Figure 3.1: An illustration of direction finding problem (DF1)

First, we consider the direction finding problem associated with GLP in the cutting plane form. From the master problem (M1), it is clear that the approximation of L is

$$L_{A1}(v) = \min \{ f(x_i) + v[g(x_i)-b]: i=0,\ldots,k-1 \}.$$

Then, Hearn and Lawphongpanich (1987) showed that u_{k+1}, the solution to the (k+1)-st master problem, solves the following problem derived from (DF1):

$$\text{DF2:} \quad \max_{v} \ [g(x_k)-b](v-u_k)$$

$$\text{s.t.} \quad w_{k+1} \le L_{A1}(v),$$

$$v \ge 0,$$

where w_{k+1}, the optimal objective function value of the (k+1)-st master problem, replaces α. Define $X(u_k)$ as the set of solutions to (S1). When $X(u_k)$ is a singleton, L is differentiable at u_k, and $[g(x_k)-b]$ is the only choice of 'subgradient' for the objective function of (DF2). However, when $X(u_k)$ is not a singleton, L is nondifferentiable and the choice of objective functions becomes infinite. Moreover, an arbitrarily chosen subgradient could result in (DF2) generating a nonascent direction as shown in Hearn and Lawphongpanich (1987). Below, we consider a second direction finding problem in which $[g(x_k)-b]$ is replaced with a set that includes the subdifferential of L at u_k.

Assume that X can be represented by a finite discrete set $\{y_1,\ldots,y_T\}$. For example, when (P1) is a bounded integer program, then X is itself a finite set of vectors with integral elements.

Then, $X(u_k)$ is a finite discrete set. For the discussion below, define for any $u \geq 0$

$$X_\varepsilon(u) = \{ y_j: f(y_j)+u(g(y_j)-b) \leq L(u) + \varepsilon,\ y_j \in X \},$$

that is, $X_\varepsilon(u)$ is the set of all ε-approximate solutions of (S1). When $\varepsilon = 0$, $X_0(u)$ is simply the set of solutions to (S1) which is previously denoted as $X(u)$. Since L is concave, the ε-subdifferential of L is well defined and can be written as follows (see, e.g., Clark, 1975; Kiwiel, 1985; Lemarechel, 1980; and Zowe, 1987):

$$\delta_\varepsilon L(u) = \{ h: L(v) \leq L(u) + h(v-u) + \varepsilon,\ \forall\ v \geq 0 \}$$

Similarly, when $\varepsilon = 0$, $\delta_0 L(u) = \delta L(u)$, the subdifferential of L at u, and $\delta L(u)$ is also equivalent to conv{ $[g(y_j)-b]$: $y_j \in X(u)$}, where conv{$\cdot$} denotes the convex hull of a set. Below, Theorem 3.1 relates $X_\varepsilon(u)$ to the ε-subdifferential of $L(u)$.

<u>Theorem 3.1</u>: Conv{ $[g(y_j)-b]$: $y_j \in X_\varepsilon(u)$ } is a subset of $\delta_\varepsilon L(u)$.

<u>Proof</u>: For every $v \geq 0$ and for each $y_j \in X_\varepsilon(u)$

$$\begin{aligned} L(u) + [g(y_j)-b](v-u) &= L(u) - [g(y_j)-b]u + [g(y_j)-b]v \\ &\geq f(y_j) - \varepsilon + [g(y_j)-b]v \\ &\geq L(v) - \varepsilon, \text{ or} \end{aligned}$$

$$L(u) + [g(y_j)-b](v-u) + \varepsilon \geq L(v), \tag{3.1}$$

where the first inequality follows from the definition of $X_\varepsilon(u)$ and the second inequality from the definition of L. Since equation (3.1) holds for all $y_j \in X_\epsilon(u)$, it must hold for all convex combinations of $y_j \in X_\epsilon(u)$, and the theorem follows. ∎

Given the above definitions, (DF1) can be modified as follows:

DF3: $$\max \min \{ [g(y_j)-b](v-u_k) : y_j \in X_\epsilon(u_k) \}$$
$$\text{s.t.} \quad L(u_k) \le L_{A2}(v)$$
$$v \ge 0,$$

where

$$L_{A2} = \min \{ f(x) + v[g(x)-b] : x \in M^{k-1} \cup \{x_0\} \}, \text{ and}$$

$$M^{k-1} = \text{ any subset of } \bigcup_{i=1}^{k-1} X_\epsilon(u_i).$$

If $X(u_k)$ replaces $X_\epsilon(u_k)$ in the objective function of (DF3), the 'min' part in the objective is the expression for the directional derivative of L at u_k in the direction $(v-u_k)$. Note that (DF3) can be written in the cutting plane format as follows:

DF4: $$\max \quad w$$
$$\text{s.t.} \quad w \le [g(y_j)-b](v-u_k) \qquad \forall\, y_j \in X_\epsilon(u_k)$$
$$L(u_k) \le f(x) + v[g(x)-b] \qquad \forall\, x \in M^{k-1} \cup \{x_0\}$$
$$v \ge 0.$$

Under the assumption that X is a finite discrete set, L(u) can be equivalently written as a minimum of a finite number of linear functions. Then, (DF3) and (DF4) resemble the direction finding problem of an algorithm proposed by Demyanov (see, Demyanov and Malozemov, 1974, and Lemarechel, 1980) in that they require a full knowledge of the ε-subdifferential. However, Demyanov's direction finding problem is nonlinear whereas (DF4) is linear.

As stated in (DF3), M^{k-1} can be any subset of the union of the previously calculated ε-subdifferentials. Theoretically, M^{k-1} can be an empty set and the algorithm to be presented in the next section would still converge. However, taking M^{k-1} as an empty set means that L(v) is approximated by only one hyperplane defined by x_o. In practice, this may not yield good directions. Thus, the choice of the set M^{k-1} should allow a good approximation of L(v) and, in turn, its contours in the neighborhood of u_k.

4. An Ascent Algorithm and its Convergence

The algorithm below is a modification of the cutting plane algorithm in which we replace the master problem (M1) with the direction finding problem (DF3) [or, equivalently, (DF4)] and add a line search step.

An Ascent Algorithm

Step 0. Let $x_o \in X$ satify $g(x_o) < b$. Select $u_1 \geq 0$ and compute $X_\varepsilon(u_1)$. Set $M^o = \emptyset$ and $k = 1$.

Step 1. (Direction Finding) Solve problem (DF3) (or, equivalently, (DF4)) and let w_k denote the value of the optimal objective function, and v_k denote the solution. If $w_k \leq 0$, stop and u_k is an optimal solution. Otherwise, go to Step 2.

Step 2. (Line search) Solve

$$L(u_k + t_k(v_k-u_k)) = \max_t \{ L(u_k + t(v_k-u_k)) : 0 \leq t \leq 1 \}$$

and set $u_{k+1} = u_k + t_k(v_k-u_k)$. Go to Step 3.

Step 3. (Evaluate $L(u_{k+1})$) Solve the subproblem (S1) and construct the set $X_\varepsilon(u_{k+1})$. Set $k = k+1$, and go to Step 1.

In Step 0, the point x_o satifies the Slater constraint qualification, and by construction x_o is included in the approximating function L_{A2}. This prevents v_k from being unbounded since $v[g(x_o)-b] \longrightarrow -\infty$ as if any component of v goes to ∞. For the convergence analysis below, it is assumed that the set {u: α

$\leq L(u)\}$ is bounded for all α. Thus, u_k is bounded for all k as well.

The first three theorems justify the direction finding problem (DF3).

<u>Theorem 4.1</u>: u_k is a solution to (D1) if and only if u_k solves (DF3).

<u>Proof</u>: For convenience, assume that

1) $X(u_k) = \{y_1, \ldots, y_p\}$,

2) $X_\varepsilon(u_k) = \{y_1, \ldots, y_p, y_{p+1}, \ldots, y_q\}$, and

3) $M^{k-1} = \{x_1, \ldots, x_c\}$.

where $p \leq q$ and $q \leq T$.

Let u_k be a solution to (D1). Then, there exists a (π, α) satisfying the following Karush-Kuhn-Tucker (KKT) conditions:

$$\sum_{j=1}^{p} \pi_j [g(y_j) - b] + \sum_{r=1}^{m} \alpha_r e_r = 0$$

$$\sum_{j=1}^{p} \pi_i = 1$$

$$\alpha_r [u_k]_r = 0, \quad r = 1, \ldots, m$$

$$\pi \text{ and } \alpha \geq 0.$$

where e_r is the r-th unit vector in R^m and $[z]_r$ denotes the r-th component of the vector z. Given this pair of multipliers, (π, α), define the triplet (π', β', α') as follows:

$$\pi'_j = \pi_j, \quad j = 1, \ldots, p$$

$$\pi'_j = 0, \quad j = p+1,\dots,q$$

$$\beta'_i = 0, \quad i = 1,\dots,c$$

$$\alpha'_r = \alpha_r, \quad r = 1,\dots,m.$$

Then, (π',β',α') satisfies the following KKT conditions for (DF3) at the point u_k:

$$\sum_{j=1}^{q} \pi'_j[g(y_j)-b] + \sum_{i=1}^{c} \beta'_i[g(x_i)-b] + \sum_{r=1}^{m} \alpha'_r e_r = 0$$

$$\sum_{j=1}^{q} \pi'_j = 1$$

$$\beta'_i[g(x_i)-b] = 0$$

$$\alpha'_r[u_k]_r = 0$$

$$\pi',\ \beta', \text{and } \alpha' \geq 0.$$

Since u_k is feasible to (DF3), u_k must be optimal to (DF3) as well.

Assume that u_k solves (DF3). Then, there exists a (π',β',α') satisfying the above KKT conditions for (DF3). Since π' and β' are nonnegative and $\Sigma_j\pi'_j = 1$, we can define

$$a = \Sigma_j\pi'_j + \Sigma_i\beta'_i$$

$$\pi^*_j = \pi'_j/a, \quad j = 1,\dots,p,p+1,\dots,q$$

$$\beta^*_i = \beta'_i/a, \quad i = 1,\dots,s$$

$$\alpha^*_r = \alpha'_r/a, \quad r = 1,\dots,m.$$

Then, (π^*,β^*,α^*) is an optimal set of multipliers at the point $(z,v) = (L(u_k),u_k)$ for the following problem which is related to (DF4):

DF5: $\max \quad z$

$$\text{s.t.} \quad z \le f(y_j) + [g(y_j)-b]v, \quad j = 1,\ldots,p,p+1,\ldots q$$

$$z \le f(x_i) + [g(x_i)-b]v, \quad i = 1,\ldots,s$$

$$v \ge 0,$$

that is, $(L(u_k),u_k)$ solves DF5. However, note that (D1) can also be written as:

D2: $\max \quad z$

$$\text{s.t.} \quad z \le L(v)$$

$$v \ge 0.$$

Then, (DF5) is a relaxation of (D2). Since $(L(u_k),u_k)$ solves (DF5) and is feasible (D2), it must be optimal to (D2).■

Theorem 4.2: If u_k does not solve (D1), then

$$\min\{ [g(y_j)-b](v_k-u_k): y_j \in X_e(u_k) \} > 0,$$

where v_k solves (DF3).

Proof: By Theorem 4.1, u_k does not solve (D1) implies that u_k is not a solution to (DF3). However, u_k is still feasible to (DF3), so the following inequality must hold:

$$\min\{ [g(y_j)-b](v_k-u_k): y_j \in X_e(u_k) \}$$

$$= \max_{v \text{ feasible}} \{ \min\{ [g(y_j)-b](v-u_k): y_j \in X_e(u_k) \}\}$$

$$> \min\{ [g(y_j)-b](u_k-u_k): y_j \in X_e(u_k) \} = 0.$$■

It should be noted that the above two results also hold when $L_{A2}(v)$ is replaced by the actual Lagrangian function L in (DF3).

The following series of lemmas and a thecrem demonstrate that the ascent algorithm above converges to an optimal solution of (D1). Furthermore, they are similar to the standard argument for establishing convergence of feasible direction algorithms in nonlinear programming.

<u>Lemma 4.3</u>: There exists an $\varepsilon > 0$ such that if $u_k \longrightarrow u*$, then

$X_\varepsilon(u_k) = X(u*)$, for k sufficiently large.

<u>Proof</u>: Let $X(u*) = \{y_1, \ldots, y_p\}$. Define

i) $h_1(u) = \max \{f(y_j) + u[g(y_j)-b]: j = 1,\ldots,p\}$

ii) $h_2(u) = \min \{f(y_j) + u[g(y_j)-b]: j = p+1,\ldots,T\}$

iii) $\alpha = [h_1(u*)+h_2(u*)]/2$

At u*, the following hold:

A) $h_1(u*) = L(u*) = f(y_j) + u*[g(y_j)-b], \quad j = 1,\ldots,p$

B) $h_1(u*) < \alpha < h_2(u*)$

C) $\sigma = [\alpha - h_1(u*)] > 0$

From the continuity of $h_1(u)$, $h_2(u)$, and $L(u)$, there must exist an integer K sufficiently large so that for every $k \geq K$

D) $|h_1(u_k) - h_1(u*)| < \sigma/2$

E) $|L(u_k) - L(u*)| < \sigma/4$

F) for $j = 1,\ldots,p$

$$|f(y_j)+u_k[g(y_j)-b] - L(u*)|$$
$$= |f(y_j)+u_k[g(y_j)-b] - f(y_j)-u*[g(y_j)-b]|$$
$$< \sigma/4.$$

G) $h_2(u_k) > \alpha$.

From (E) and (F), we have that for $j = 1,...,p$

$$|f(y_j)+u_k[g(y_j)-b] - L(u_k)|$$

$$\leq |f(y_j)+u_k[g(y_j)-b] - L(u*)| + |L(u*) - L(u_k)|$$

$$\leq \sigma/2,$$

and since $L(u_k) \leq f(y_j) + u_k[g(y_j)-b]$, $\forall$ j, it follows that

$$|f(y_j)+u_k[g(y_j)-b] - L(u_k)| = f(y_j)+u_k[g(y_j)-b] - L(u_k) \leq \sigma/2, \text{ or}$$

$$f(y_j)+u_k[g(y_j)-b] \leq L(u_k)+\sigma/2.$$

Thus, for any ε between $(1/2)\sigma$ and $(3/4)\sigma$,

$$f(y_j) + u_k[g(y_j)-b] \leq L(u_k) + \varepsilon, \quad \text{for } j = 1,...,p, \tag{4.1}$$

that is, y_j is an element of $X_\varepsilon(u_k)$ for $j = 1,...,p$. Moreover, from (E),

$$\begin{aligned} L(u_k) &\leq L(u*) + \sigma/4 \\ L(u_k) + \varepsilon &\leq L(u*) + \sigma/4 + \varepsilon \\ &\leq L(u*) + \sigma \\ &= \alpha \\ &< f(y_j)+u_k[g(y_j)-b], \quad \text{for } j = p+1,...,T. \end{aligned} \tag{4.2}$$

where the third inequality follows the above selection of ε, the equality from (C), and the last inequality from (G) and the definition of $h_2(u)$. This means that y_j does not belong to $X_\varepsilon(u_k)$, for $j = p+1,...,T$. Therefore, $X_\varepsilon(u_k) = X(u*)$. ∎

Lemma 4.4: For a given direction d,

$$\min\{\, g^{\hat{}}d:\ g^{\hat{}} \in \operatorname{conv}([g(y_j)-b]:\ j = 1,\ldots,p)\,\}$$

$$= \min\{\, [g(y_j)-b]d:\ j = 1,\ldots,p\,\}$$

Proof: The result follows from the fact that the problem on the left hand side can be stated as a linear program

$$\begin{aligned} \min \quad & \sum_{j=1}^{p} ([g(y_j)-b]d)\pi_j \\ \text{s.t.} \quad & \sum_{j=1}^{p} \pi_j = 1 \\ & \pi_j \ge 0, \quad j = 1,\ldots,p, \end{aligned}$$

which always yields an extreme point solution. ∎

Lemma 4.5: If $X_e(u_k) = \{y_1,\ldots,y_p\}$ and $\min\{\, [g(y_j)-b]d:\ j = 1,\ldots,p\,\} > 0$, then there exists a $\tau > 0$ such that

$$L(u_k+\sigma d) \ge L(u_k)+\sigma\min\{\, [g(y_j)-b]d:\ j = 1,\ldots,p\,\}$$

for all $0 \le \sigma \le \tau$.

Proof: Assume without loss of generality that

$$0 < [g(y_1)-b]d = \min\{\, [g(y_j)-b]d:\ j = 1,\ldots,p\,\}$$

Define

$$\tau = \min\left[\frac{\{f(y_j)+u[g(y_j)-b] - L(u_k)\}}{[g(y_1) - g(y_j)]d},\ \text{for all } j \text{ such that } g(y_1)d > g(y_j)d \ \&\ j \notin X_e(u_k) \right]$$

and observe that $\tau > 0$ since

$$f(y_j) + u[g(y_j)-b] > L(u_k) \qquad \text{for } j \notin X_e(u_k),$$

and $[g(y_i)-g(y_j)]d > 0$ by construction. Thus, for any $0 \le \sigma \le \tau$ and every $j \notin X_\epsilon(u_k)$ and $g(y_i)d > g(y_j)d$

$$\sigma \le \frac{\{f(y_j) + u_k[g(y_j)-b] - L(u_k)}{[g(y_i) - g(y_j)]d}$$

$$\sigma[g(y_i)-b - g(y_j)+b]d \le f(y_j) + u_k[g(y_j)-b] - L(u_k)$$

$$L(u_k) + \sigma[g(y_i)-b]d \le f(y_j) + (u_k+\sigma d)[g(u_j)-b] \qquad (4.3)$$

However, for $j \notin X_\epsilon(u_k)$ and $0 < [g(y_i)-b]d \le [g(y_j)-b]d$

$$L(u_k) < f(y_j) + u_k[g(y_j)-b]$$

$$L(u_k) + \sigma[g(y_i)-b]d < f(y_j) + (u_k+\sigma d)[g(y_j)-b] \qquad (4.4)$$

and for $j \in X_\epsilon(u_k)$

$$L(u_k) \le f(y_j) + u_k[g(y_j)-b]$$

$$L(u_k) + \sigma[g(y_i)-b]d \le f(y_j) + (u_k+\sigma d)[g(y_j)-b]. \qquad (4.5)$$

Combining (4.3), (4.4) and (4.5), we have that for $0 \le \sigma \le \tau$

$$L(u_k) + \sigma[g(y_i)-b]d \le \min\{ f(y_j)+(u_k+\sigma d)[g(y_j)-b]: j=1,...,T\}$$
$$= L(u_k+\sigma d). \blacksquare$$

<u>Lemma 4.6</u>: Assume that ϵ is chosen as in Lemma 4.3 and the algorithm generates a sequence $\{u_k\}$. Then, there cannot be a subsequence $\{u_k\}$, $k \in \Omega$, with the following properties:

i) $u_k \longrightarrow u_\infty, \quad k \in \Omega,$

ii) $v_k \longrightarrow v_\infty, \quad k \in \Omega,$ and

iii) $\min\{ g^*(v_\infty-u_\infty): g^* \in \mathrm{conv}([g(y)-b]: y \in X(u_\infty)) \} > 0,$

where Ω is a subset of $\{1,2,3,...\}$.

Proof: Assume that $X(u_\infty) = \{1,\ldots,p\}$. By (iii) and Lemma 4.4,

$$\min\{ g^*(v_\infty-u_\infty): g^* \varepsilon \operatorname{conv}([g(y)-b]: y \varepsilon X(u_\infty)) \}$$
$$= \min \{ [g(y_j)-b](v_\infty-u_\infty): j = 1,\ldots,p \} = \beta > 0.$$

From Lemma 4.3, there must exist a K^1 sufficiently large such that $X_\varepsilon(u_k) = \{1,\ldots,p\}$ for some $k \geq K^1$ and $k \varepsilon \Omega$. Since the set $\{1,\ldots p\}$ is finite, there must exist a index $j*$ such that

$$j* = \arg\min\{ [g(y_j)-b](v_k-u_k): j = 1,\ldots,p \} \qquad (4.6)$$

infinitely often. Define Ω^1 to be the subset of Ω for which $j*$ is the index which yields the minimum value for the right hand side of (4.6). For convenience, we assume that $j* = 1$. Then, we have that

$$\lim_{k \varepsilon \Omega^1} [g(y_1)-b](v_k-u_k) = [g(y_1)-b](v_\infty-u_\infty) = \beta > 0,$$

and it follows that there exists $K^2 \geq K^1$ such that

$$[g(y_1)-b](v_k-u_k) > \beta/2 \quad \text{for } k \geq K^2 \text{ and } k \varepsilon \Omega^1. \qquad (4.7)$$

Thus, at u_k the direction (v_k-u_k) is an ascent direction. Moreover, since u_{k+1} maximizes $L(u)$ along the direction (v_k-u_k),

$$L(u_{k+1}) \geq L(u_k + \sigma(v_k-u_k)) \quad \text{for } 0 < \sigma < \tau \text{ and } k \varepsilon \Omega^1$$
$$\geq L(u_k) + \sigma[g(y_1)-b](v_k-u_k) \quad \text{for } 0 < \sigma < \tau \text{ and } k \varepsilon \Omega^1$$
$$\geq L(u_k) + (\sigma\beta)/2 \quad \text{for } 0 < \sigma < \tau \text{ and } k \varepsilon \Omega^1, \qquad (4.8)$$

where τ is as defined in Lemma 4.5 and the last two inequalities follow from Lemma 4.5 and equ.(4.7). Letting k approach infinity on the subsequence Ω^1, (4.8) yields

$$L(u_\infty) \geq L(u_\infty) + (\sigma\beta)/2$$

which is a contradiction since both σ and β are positive.∎

<u>Theorem 4.7</u>: Assume that ε is chosen as in Lemma 4.3. If the algorithm generates a sequence $\{u_k\}$, then it must converge to an optimal solution of (D1).

<u>Proof</u>: Assume that the algorithm generates an infinite sequence of points and none of which is a solution to (D1). Since $\{u_k\}$ lies in a bounded region, there must exist a convergent subsequence, i.e., $u_k \longrightarrow u_\infty$, for $k \in \Omega$, where Ω is a subset of $\{1,2,3,...\}$ and u_∞ does not solve (D1).

Assume that $X(u_\infty) = \{1,...,p\}$. By Lemma 4.3, $X_\varepsilon(u_k) = X(u_\infty)$ for k sufficiently large, i.e., $k \geq K_1$, thus $X_\varepsilon(u_\infty) = X(u_\infty)$. Since u_∞ is not a solution to (D1),

$$\begin{aligned} 0 < \beta_\infty = \max_{z}\ & \min_{j}\{ [g(y_j)-b](z-u_\infty) : j = 1,...,p \} \\ \text{s.t.}\ & L(u_\infty) \leq L(z) \\ & z \geq 0. \end{aligned}$$

For $k \geq K_1$, define

$$\text{DF6:} \quad \begin{aligned} \beta_k = \max_{z}\ & \min_{j}\{ [g(y_j)-b](z-u_k) : j = 1,...,p \} \\ \text{s.t.}\ & L(u_k) \leq L(z) \\ & z \geq 0. \end{aligned}$$

Note that β_k varies continuously with u_k. Thus, $\beta_k \longrightarrow \beta_\infty$ for $k \in \Omega$ and $k \geq K_1$.

Now, let v_k be a solution to (DF3) defined at u_k. Then, by construction, v_k is bounded and there must exist a subset, Ω^1, of Ω such that $v_k \longrightarrow v_\infty$ on Ω^1. Moreover, since (DF3) is a relaxation of (DF6),

$$\beta_k \le \min\{ [g(y_j)-b](v_k-u_k) : j = 1,\ldots,p \}, \tag{4.9}$$

for $k \ge K_1$ and $k \in \Omega^1$. From the finiteness of of the set $\{1,\ldots,p\}$, there must exist an integer j^* such that

$$j^* = \arg\min\{ [g(y_j)-b](v_k-u_k) : j = 1,\ldots,p \} \tag{4.10}$$

infinitely often. Let Ω^2 be a subset of Ω^1 for which j^* is the index which minimizes the right hand side of (4.10) and for convenience assume that $j^* = 1$. Then, combining (4.9) and (4.10) gives

$$[g(y_1)-b](v_k-u_k) \ge \beta_k \qquad \text{for } k \ge K_1 \text{ and } \Omega^2.$$

Taking the limit on both side with respect to Ω^2, we have that

$$\lim\, [g(y_1)-b](v_k-u_k) = [g(y_1)-b](v_\infty-u_\infty) \ge \beta_\infty > 0$$

which contradicts Lemma 4.6. ∎

Therefore, if the algorithm terminates after a finite number of iterations, Theorem 4.1 and 4.2 guarantee that u_k solves (D1). Otherwise, the algorithm generates an infinite sequence which, by Theorem 4.7, converges to an optimal solution of (D1). Also, it is interesting to note that although the algorithm uses ε-subdifferential in calculating its ascent directions an exact optimal solution can be obtained by choosing

ε correctly. In general, one expects algorithms using an ε-subdifferential to produce ε-optimal solutions, possibly in a finite number of iterations. We consider this type of algorithms in a separate study to appear later.

References

Clark, F. H. (1975), "Generalized Gradients and Applications," Transaction of American Mathematical Society, 205, pp. 247-262.

Dantzig, G. B. (1963), Linear Programming and Extensions, Princton University Press, Princeton, NJ.

Dantzig, G. B., and Wolfe, P. (1960), "Decomposition Principle for Linear Programs," Operations Research 8, pp. 101-111.

Dantzig, G. B., and Wolfe, P. (1961), "The Decomposition Algorithm for Linear Programming," Econometrica 29, pp. 767-778.

Demyanov, V. F. and Malozemov, V. N. (1974), Introduction to Minimax, John Wiley & Sons, New York, NY.

Geoffrion, A. M. (1970), "Elements of Large-Scale Mathematical Programming," Management Science 16, pp. 652-691.

Hearn, D. W. and Lawphongpanich, S. (1987), "Lagrangian Dual Ascent by Generalized Linear Programming," Research Report No. 87-14, Deparment of Industrial and Systems Engineering, University of Florida, Gainesville, Florida 32611.

Lemarechel, C. (1980), "Nondifferential Optimization," Nonlinear Optimization, Theory and Algorithms, Dixon, Spedicato and Szegö (EDs.), Birkhäuser, Boston, MA., pp. 151-199.

Kiwiel, K. C. (1985), Methods of Descent for Nondifferentiable Optimization, Lecture Notes in Mathematics 1133, Springer-Verlag, New York, NY.

Zangwill, W. I. (1969), Nonlinear Programming: A Unified Approach., Prentice Hall, Englewood Cliffs, NJ.

Zowe, J. (1987), "Nondifferentiable Optimization," Computational Mathematical Programming, NATO ASI Series Vol. F15, Schittkowski (Ed.), Springer-Verlag, New York, NY, pp. 323-356.

DISTRIBUTION LIST

	NO. OF COPIES
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5000	2
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Office of Research Administration (Code 012) Naval Postgraduate School Monterey, CA 93943-5000	1
Center for Naval Analyses 4401 Ford Ave. Alexandria, VA 22311	1
Library (Code 55) Naval Postgraduate School Monterey, CA 93943-5000	1
Operations Research Center, Rm E40-164 Massachusetts Institute of Technology Attn: R. C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Koh Peng Kong OA Branch, DSO Ministry of Defense Blk 29 Middlesex Road SINGAPORE 1024	1
Arthur P. Hurter, Jr. Professor and Chairman Dept of Industrial Engineering and Management Sciences Northwestern University Evanston, IL 60201-9990	1
Institute for Defense Analysis 1800 North Beauregard Alexandria, VA 22311	1
Chief of Naval Research 800 N. Quincy St. Arlington, VA 22217-5000	1

DUDLEY KNOX LIBRARY
3 2768 00329687 2